Reconstructed Intentions in Collaborative Problem Solving Dialogues

Barbara Di Eugenio* **Pamela W. Jordan**†

Richmond H. Thomason‡ **Johanna D. Moore**§

University of Pittsburgh
Pittsburgh, PA 15260
URL: http://www.isp.pitt.edu/~intgen/

Introduction

In this paper, we would like to address the question posed in the Call for Papers about whether the *speech act* is a useful intermediary concept (e.g., for representing intentions), or whether a direct *context-change* model is more appropriate to the task. In our opinion, the concept of *speech act* is a useful intermediary between surface form and intentions, and we do not advocate doing away with it; however, we also think that speech act recognition is difficult, and that both humans and computers can make use of additional evidence that indicates what is, or has to be added, to the common ground. While providing evidence for our position, we will discuss these issues in the context of recognizing *proposes* and *accepts*.

In general, speech act recognition plays a prominent role in dialogue understanding, in traditional approaches that infer a plan using plan construction operators (Perrault & Allen 1980), (Litman & Allen 1990), (Lambert & Carberry 1991; 1992), and in more recent techniques relying on statistical correlations or finite state machines (Reithinger & Maier 1995; Qu *et al.* 1997). Both approaches recognize surface speech acts, using surface form and information provided by the discourse context and the discourse operators, or by a finite state approximation of the planning information.

These approaches assume that it is (relatively) simple to recognize speech acts, and that speech acts are a required element of the analysis, corresponding closely to the speaker's intentions. In this paper, we provide evidence that:

(i) It can be very difficult for humans to reliably recognize speech acts, and

(ii) In some cases the association of a speech act with an utterance can reconstruct an intention far more determinate than anything the speaker entertained at the time of utterance.

In such cases, the problem solving context, together with certain surface cues that are not usually associated with speech act recognition, prove to be more predictive than the methods that are typically used in speech act recognition.

Reconstructed proposals

We will illustrate these problems with reference to RECONSTRUCTED proposals, in the context of our collaborative problem solving corpus. Our corpus was collected in experiments in which subjects are asked to buy furniture for two rooms in a house.

An agent engaged in collaborative problem solving must recognize when an agreement has been reached on sub-parts of the problem. One plausible approach to recognizing that an agreement has been reached is to recognize its components; propose and accept.[1] However this approach is not always reliable.

(i) A: I do not have a sofa for a better price but, i do have a lamp-floor, blue (250). i have a green table (200) and four chairs for (75) a piece.

(ii) B: ... the lamp and table sound good,

In sentence (i), agent A informs B that he has a floor lamp and in (ii) agent B takes this as a proposal to buy the lamp. A may well not intend to propose the floor lamp, but because B knows that she does not have a better alternative, she opportunistically treats (i) as a propose and accepts it. If A doesn't object, the conversation continues as if A had indeed uttered a propose; and it is immaterial whether A definitely intended to propose the lamp when (i) was uttered.[2]

This is a matter of vague rather than of ambiguous intentions. So the interpretation can't just be a matter of choosing the most likely of several alternatives.

*Learning Research & Development Center; dieugeni@cs.pitt.edu

†Intelligent Systems Program; jordan@isp.pitt.edu

‡Intelligent Systems Program; thomason@isp.pitt.edu

§Computer Science Department, and Learning Research & Development Center; jmoore@cs.pitt.edu

[1] Compare the approach of (Lochbaum 1994, p. 28), which is essentially componential, in that the recognition of an intention to perform a (possibly joint) act depends on the prior recognition of a desire to perform that act.

[2] See (Fox 1987) for background on this sort of retrospective reinterpretation.

Our corpus makes it clear that reconstructed proposals are a natural and common way of doing business in some genres. Therefore, accounting for them is an important part of understanding the discourses in which they occur, despite the empirical problems of dealing with them.

Modeling the role of context

We believe that the problem of reconstructed proposals is connected with the the role of context in generation and interpretation. It has been known for a long while that speech acts are highly sensitive to features of the discourse context (Strawson 1964). For instance, there are contexts in which an utterance like *I guess I can study tomorrow* is an acceptance of an offer. In these contexts, the main point of the utterance will be missed if the role of context is ignored.[3] Even if speech acts are not recognized, context plays a fundamental role because it affects the effects a certain utterance has.

In our larger project, we are seeking to model interpretive reasoning of this sort using the following ingredients:

(i) A reasoning process that is able to choose a likely and preferred interpretation from among several alternatives.
(ii) A model of the context in which the interpretation occurs.
(iii) A mechanism for maintaining contexts, that allows them to be updated in the light of conversational information and common knowledge.
(iv) A general way of allowing these contexts to influence the interpretive process. In particular, contexts should influence the preferences that are assigned to interpretations.

Our approach (Thomason & Moore 1995; Thomason & Hobbs 1997) uses a form of weighted abduction as the reasoning mechanism (Stickel 1988; Hobbs *et al.* 1993) with contexts modeled as modal-like operators (ingredient ii). This allows contexts to be represented explicitly and reasoned about.[4] The abductive interpretation of an utterance will use a modal operator to limit the rules that are accessible and adjust the weights of assumptions. We can model contexts that favor the interpretation of an utterance like *I don't have any tables* as an accept. Similarly (since we also take an abductive approach to generation), we can also favor the generation of this utterance as a way of accepting in these contexts.

As described in (Thomason & Hobbs 1997), when a speaker desires a change in the conversational record, she finds a speech act that will bring about this change and then finds an utterance that embodies the speech act. Recognizing a desired change in the conversational record works in the reverse. However, to interpret the earlier example with the lamp and the table as a reconstructed proposal, regardless of whether a propose speech act was intended by the speaker, we rely on the context to select a modified set of axioms that will allow an alternate way of recognizing a propose speech act. This alternative means of recognition circumvents the need for an explicit utterance that will embody the propose speech act. The context change affects the speech acts but not the other reasoning implemented by the axioms.

The question then is how to recognize the appropriate context. We do this by mapping from the *informal* context, which includes such components as the domain reasoning situation and the discourse history, to a modal operator that represents the formal context. Our approach for finding the appropriate mapping is to look for guidance from correlations between corpus and domain features and a particular modal operator.

Our working hypothesis for modeling proposals then is that the choice of context (ingredient iii) is partly influenced, for both the speaker S and the hearer H, by the domain reasoning situation. In particular, if the suitable courses of action are highly limited, this will make a previous utterance more likely to be treated as a propose during response generation (ingredient i). This means that a context that allows an accept without an explicit propose can be selected during response generation (but as we will see from the corpus analysis, the response must also use an appropriate surface form and coreference relation). During interpretation of this type of response, the suitable courses of action must be limited as an effect of the utterance, and this in combination with coreference and surface form will make the utterance more likely to be interpreted as an accept (ingredient i). In both generation and interpretation, the additional predictive power of the domain reasoning selects the context which in turn adjusts the weights of axioms and assumptions (ingredient iv).

The remainder of the paper is organized as follows. First we describe our domain and then we discuss our evidence for the difficulty of reliably recognizing speech acts. Finally, in the last two sections we return to the problem of reconstructed proposals and show the preliminary results of a test of our working hypothesis.

The Collaborative Problem Solving Corpus

The subjects in our collected conversations are equal in status: they were both briefed on the domain knowledge needed for problem solving and neither is an expert at this task. The task is to buy furniture for the living and dining rooms of a house. (The task is based on those in (Walker 1993; Whittaker, Geelhoed, & Robinson 1993)). Each subject is given a separate budget and inventory of furniture that lists the quantities, colors, and prices for each available item. By sharing this information during their conversation, the

[3] Green and Carberry make a similar point about indirect answers in (Green & Carberry 1994).

[4] See (McCarthy & Buvač 1995) for ideas about the formalization of context and contextual reasoning.

subjects can combine their budgets and can select furniture from each other's inventories. The problem is collaborative in that all decisions have to be consensual; funds are shared and purchasing decisions are joint. Subjects are asked to maintain private graphical representations of their discussions and incremental agreements. We use this private information as partial evidence of what S's utterance meant and what H understood.

The subjects' main goal is to negotiate the purchases; the items of highest priority are a sofa for the living room and a table and four chairs for the dining room. The subjects also have specific secondary goals which further complicate the problem solving task. Subjects are instructed to try to meet as many of these goals as possible. The secondary goals are: 1) Match colors within a room, 2) Buy as much furniture as you can, 3) Spend all your money.

Identifying Speech Acts

We argue that it can be difficult for humans to recognize speech acts, by looking at the relative difficulty of achieving a reliable coding for speech acts compared to other features of problem solving dialogues. Although relative difficulty is merely anecdotal at this point, if one can achieve a reliable coding for other features while still struggling to do so for speech acts, it seems worthwhile to find out if a combination of these other features might suffice.

Coding Schema

Our tagging scheme is an augmented version of the Discourse Resource Initiative (DRI) one,[5] which is based on the familiar taxonomy of (Searle 1975). We devised our coding schema with two goals in mind: to conform as much as possible with the standards for mark-up being developed within DRI, and to represent the important features of the discourse that we collected. As with DRI, we tag aspects that are inherent to the utterance itself, and that encode the relationship of the utterance(s) to the preceding utterance(s). The tags of interest for this paper are:

(i) Utterance level tags: topic and illocutionary act.

(ii) Inter-utterance tags: relational tags (in particular, response-to) and coreference tags.

Topic tags[6] capture the meaning of the utterance by encoding what is relevant to the problem solving task in terms of furniture items or money: e.g., S may either state that he has a particular item (*I have a blue sofa for $300.*); discuss selecting a particular item (*shall we buy the two red chairs*); elaborate the description of an item that has already been introduced (*my red chairs are $100 each*); express an evaluation with respect to a specific furniture item (*the chairs seem expensive*); or discuss the budget (*I have $300.*).

Illocutionary-Act tags capture the intention behind the utterance and characterize at an abstract level S's main intention. At the highest level, the choices are *Inform, Directive, Commissive,* and *Conventional.*

An *Inform* utterance is intended to get H to believe something while a *Directive* is intended to get H to do something. Directives are further subdivided into: *Request-Action*, as in *Buy the chairs*; *Request-Info*, where the action requested is that H provides the desired information—many questions will fall under this category, e.g. *What do we have left if anything?*; and *Suggest*, which is a *Request-Action* that is conditional on agreement with H, as in *How about buying those two chairs.*

The primary purpose of a *Commissive* is to commit S (in varying degrees of strength) to some course of action. *Commissives* are subdivided into *Promise* and *Offer*. This distinction reflects the conditionality of S's commitment. S's commitment to an offer is conditional on H's agreement, so that the conversation will felicitously continue with H either accepting or rejecting. A *Promise* is not conditional in this way—or, if it depends on H's agreement, presupposes this agreement.

The distinction between *Directives* and *Commissives* is sometimes hard to draw. With a *Directive*, S asks/orders H to perform an action, while a *Commissive* constrains S's own actions as well. However in our domain, most actions are joint because they are either literally performed together by S and H, or require H's approval, even though S has initiated the action (Tuomela 1995). For example, *Let's buy the two chairs* attempts to create an agreed joint course of action by committing S to it, provided H accepts the proposal. It therefore seeks to constrain the behavior of both S and H. Taxonomies like Searle's, which classify speech acts according to whether they impose constraints on S or H, do not fit joint speech acts or even their components, which in general impose coordinated constraints on both participants (Hancher 1979). Rather than complicating the taxonomy by adding another category for joint speech acts, we arbitrarily stipulated that proposals, such as *Let's buy the two chairs*, should be tagged as *Commissive.*

Going back to the propose speech act, the reader will notice that we do not include it directly as one of our categories. This is because we expect that the definitions of a *Suggest* and an *Offer* together would cover most definitions of a propose speech act.

Relational Tags capture part of the relation between an utterance and the previous discourse: namely, an utterance or group of utterances $\{U_i\}$ can be unsolicited, or can respond to a previous utterance or

[5] The work we present in this paper is based on the first, not on the last, version of the DRI manual (Allen & Core 1996).

[6] We use the term *Topic* in a completely informal way.

	Speech-Acts Level 1	Speech-Acts Level 2	Topic	RelTags	Coref
Pair 1 (LB)	.68 (127)	.59 (127)	.81 (126)	.74 (117)	.82 (112)
Pair 2 (LP)	.51 (64)	.45 (64)	.76 (62)	.77 (60)	.74 (58)
3-way coding	.62 (35)	.60 (35)	.79 (35)	.83 (33)	.88 (32)

Table 1: Kappa values

segment.[7]

The two relational tags of interest are:

(i) *Initiate*—$\{U_i\}$ is unsolicited.

(ii) *Response-To*—$\{U_i\}$ is a response to a previous utterance or segment.

(ii.a.) *Answer*—answers a question, e.g. *No, the chairs are 100 each.*

(ii.b.) *Accept*[8]—accepts the proposal or content of its antecedent, e.g. *The chairs sound good.*

(ii.c.) *Reject*—rejects the proposal or content of its antecedent.

Coreference Relations capture how furniture items discussed in one utterance are related to those previously discussed by means of the tags *sameItem, subset* and *mut(ually)-excl(usive). SameItem* is used when an utterance is related to its antecedent via the same item or set of items. For the current utterance to be tagged as *sameItem*, it must discuss exactly the same items as the antecedent, otherwise the tag *subset* is used. The tag *mut(ually)-excl(usive)* is used when U_1 mentions a set of items $\mathcal{S}_1$, U_2 provides an alternative $\mathcal{S}_2$ to that same set of items, and $\mathcal{S}_1$ and $\mathcal{S}_2$ are mutually exclusive.

Note that an *Initiate* utterance can still be linked to the preceding discourse via coreference relations.

Analysis of the Coding Results

We coded 4 of the 12 dialogues in our corpus (approximately 1738 of 4169 words or 42% of our data). We had three coders; L,B and P. L coded all 4 dialogues, B coded dialogues 1,2 and 4 and P dialogues 3 and 4. We report here on the pair-wise coder agreement, and also the agreement on the dialogue coded by all three. Table 1 reports values for the Kappa coefficient of agreement (Carletta 1996) for five different categories—the Kappa coefficient factors out chance agreement between coders. In each cell, the first number is Kappa, and the second number is the actual size of the coded data for that particular tag.

The first column (Level 1) refers to the highest level speech act coding, e.g. *Inform, Directive, Commissive* and *Conventional*; while the second column (Level 2) also takes into account the distinctions under *Directive* and *Commissive*. The other three columns refer to *Topic, Rel(ational)Tags,* and *Coref(erence) (Relations)*. The values for RelTags are computed by taking into account the different types of *Response-to*, i.e. *Answer, Accept, Reject.*

The discourse processing community is currently using Krippendorff's scale (Krippendorff 1980) to assess Kappa's significance. The scale discounts any variable with K < .67, allows only tentative conclusions for K between .67 and .8, and allows real conclusions only for variables with K > .8. Thus, Table 1 strongly suggests that speech acts are an unreliable category. The fact that our coding schema is based on DRI makes us confident that our category distinctions and coding instructions are not arbitrary. All the same, the low coding reliability may be due to either the inadequacy of the taxonomy, a lack of clarity in the coding manual, or both. However, we can at least safely conclude that speech act recognition is a complicated enterprise.

Given the difficulties of reliably coding for speech acts, in the next sections we will consider what use we can make of the other more reliably coded features. Looking again at our working hypothesis, we would like to find correlations that will predict a regular propose context and ones that will predict an accept context. Note that in predicting an accept context, we cannot rely on the typical speech act recognition approaches mentioned in the introduction if we are to account for both regular and reconstructed proposals. Those approaches look for a preceding propose utterance as a predictor. Recall our argument that in the case of reconstructed proposals, a propose will not precede an accept.

Since we cannot reliably code for *Offers* and *Suggests* we can not search for correlations that will predict that an utterance is a propose. But since we can reliably code accepts, we can look for feature correlations that will predict that an utterance is an accept. We have the added advantage that any predictors we find for

[7] Space constraints prevent discussion of our treatment of segments.

[8] While an *Accept* is taken to be a speech act in certain frameworks, we believe it is more fruitful to see it as a dialogue act at a different level than speech acts. This viewpoint is further supported in the latest version of the Discourse Resource Initiative coding manual (Allen & Core 1997). There, the traditional taxonomy of speech acts due to Searle falls under the *forward* communicative function, while dialogue acts such as *accept* and *reject* fall under the *backward* communicative function. Note that this treatment of *Accept* is relevant to the question in the Call for Papers about what kinds of dialogue actions, other than sentence-level speech acts, occur in dialogue.

an accept will be independent of a propose speech act and should be able to account for both regular and reconstructed proposals.

The Problem Solving Model and Informal Context

Recall that in the earlier section on the role of context, our working hypothesis for modeling propose and accept depended partly on whether the suitable courses of action are highly limited. First, we will show how we classify the suitable courses of action as either highly limited (determinate) or unlimited (indeterminate).

The domain problem solving for the task (described earlier in the section on the corpus) is more readily modeled as a constraint satisfaction problem than a planning problem since the temporal ordering of *buy* actions does not affect the solution. We view the problem space as a set of variables that must have a single value or a set of values of a certain cardinality assigned to them. Since the set of possible values is not known at the outset of problem solving, the model must recognize when to treat the set of values as open, when to treat it as closed and when to reopen it.

We use the SCREAMER constraint logic programming language (Siskind & McAllester 1993) to model the problem solving. Although SCREAMER does not handle dynamic variables, we temporarily resolve this by setting up the variables and constraints anew with each utterance.

For a variable to be solved, the cardinality of the assigned value set must match the cardinality designated for the variable.[9] For example, if the *chair* variable has a designated cardinality of 3, and a shared value set of cardinality 4 (i.e. there are 4 instances of chairs that are known), but no more than 2 instances of chairs can ever be assigned without violating the budget constraint then the variable is unsolvable. For a problem solution to exist, each of the variables being considered must have a solution.

We provided the problem solving model with the constraint problem represented by each utterance as input and it gave as output the solution size for each problem. The input constraint problem represented by an utterance is limited to just the shared knowledge of S and H as an effect of the utterance, and comprises:

1. the variables being considered
2. the accumulated values for these variables
3. the current constraints

We characterized the solution size for the problem as determinate if there is one or more solutions.[10] If there is no solution to the problem or the shared value set for some variable is open we characterized it as indeterminate. A value set is open if, for example, S supplies appropriate values for a variable but does not know what H has available for this variable. No solution will be found in cases where all combinations of value assignments to some variable of interest violate some constraint. Just one unsolvable variable is all that is needed to render the problem unsolvable.

[9]Note that the cardinality designation is realized as a constraint in our implementation and is not a feature provided by SCREAMER.

[10]A threshold could be given for an upper limit on the number of solutions but the number of solutions was rarely larger than 1 or 2 since the dialogue participants typically tightly constrained the problem.

Relational Tag	Solution Size Determinate	
	Yes	No
Accept	11	1
Other	29	53

Table 2: Correlation of Acceptance and Solution Size

Relational Tag	Corefers to an utterance in a prior turn	
	Yes	No
Accept	15	2
Other	21	114

Table 3: Correlation of Acceptance and Coreference

Predicting Acceptances

In this paper, we look just at the *interpretation* component as a test of our working hypothesis. Recall that the interpretation aspect of the hypothesis involves the utterance that is meant as an accept (e.g. *the lamp and table sound good*). Informally, in this situation the hearer must recognize that his previous utterance (e.g. *I do have a lamp-floor...I have a green table..*) has been reconstructed before he can interpret the current utterance as an accept.

We first look at the predictive power of each of the coded features that we expect to play a role in recognizing the acceptance and then at how a combination of these features increases the predictive power. We will look at solution size, coreference, and topic compared to acceptances.[11]

Table 2 shows first that there is a correlation between utterances tagged as acceptances and the solution size as an effect of the utterance ($\chi^2 = 13.57, p < .001, df = 1$). Furthermore it shows that an acceptance more frequently has a determinate solution size.

Also, we see that acceptances more frequently corefer to an item in a prior utterance (Table 3) and more frequently are about getting an item or evaluating an item (Table 4).[12] While the recall for an accept is good

[11]Note that the number of features reported in each table do not necessarily agree across the tables. For example, Table 3 has 17 accepts while Table 4 has only 12. This is because each utterance was not coded for all of the features represented by these tables.

[12]In these last two tables, some expected frequencies are

Relational Tag	Topic $\in$ {getItem, eval or relate} Yes	No
Accept	12	0
Other	31	105

Table 4: Correlation of Acceptance and Topic

Relational Tag	Predictive Rule Applies Yes	No
Accept	10	2
Other	2	80

Table 5: Correlation of Acceptance and Predictive Rule

in each of the tables the precision is not. For example, if the only predictor for an accept is that the solution size is determinate, we see by looking at Table 2 that only 1 accept utterance would be missed but that 29 utterances would be falsely classified as an accept. Because the precision is bad for each feature in isolation, we expect that none of the features alone would reliably predict an accept.

Thus we combine these features into the following rule. An utterance is more likely to be an acceptance when the solution size is determinate, is linked via coreference (sameItem, subset, mut-excl) to a prior turn and either:

1. the topic is about an evaluation of an item (eval, relate) or
2. the topic is about getting an item (getItem) and the utterance linked by coreference either has a getItem topic or is determinate.

As shown in Table 5, the above rule correctly predicts a majority of the utterances labeled as acceptances and falsely predicts 2 out of 82 other utterances as acceptances.

We expect this rule to enable us to recognize that an agreement has been reached without first recognizing and assigning a propose speech act (or Suggest or Offer as in the case of our coding scheme) to the previous turn. It does this by helping us to select the appropriate context. Once the appropriate context is selected (in this case the appropriate set of axioms), abductive *interpretation* would proceed as described in the context section. Note that to apply the rule, we will need heuristics to extract the information required by the rule (i.e. topic, coreference relations, and the constraint problem definition) since the rule is applied prior to full interpretation.

too low to validly calculate χ^2. More instances of utterances are needed.

Conclusions and Future Work

We have argued that reconstructed proposals cannot be identified easily by speech act recognition and have shown that speech acts are more difficult to code than other features. We have presented a rule for predicting which utterances are acceptances based on domain context and the other more reliably coded features. This rule can be used to do context selection prior to full interpretation and allows us to uniformly solve the problem of recognizing regular and reconstructed proposals.

We are currently coding an expanded corpus using a new coding scheme (Di Eugenio, Jordan, & Pylkkänen 1997) which is based on the latest, greatly revised, Discourse Resource Initiative coding manual (Allen & Core 1997). We will use this latest coding to further test the predictive rule and will report on the new statistics at the symposium. In future empirical work, we will consider how to code and process summaries, as a means of checking agreement. We are also implementing and testing a simulation of the domain reasoning and part of the discourse generation and interpretation, including the planning and interpretation of agreements. In our simulation environment, topic and coreference relations are easy to determine but as the communication language approaches a more natural version of language, this process will be more challenging. We will also need to address the problem of extracting the variables, values and constraints needed to compute the determinacy of an utterance. This will be one of the next steps we take in implementing the simulation.

Acknowledgements

This material is based on work supported by the National Science Foundation under Grant No. IRI-9314961, "Integrated techniques for natural language generation and interpretation."

References

Allen, J., and Core, M. 1996. Draft of DAMSL: Dialog act markup in several layers. Available from http://www.cs.rochester.edu:80/research/trains/annotation under Old Version of Annotation Resources, Dialog Annotation Manual.

Allen, J., and Core, M. 1997. Draft of DAMSL: Dialog act markup in several layers. Available from http://www.cs.rochester.edu:80/research/trains/annotation under New Version of Annotation Resources, Dialog Annotation Manual.

Carletta, J. 1996. Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics* 22(2).

Di Eugenio, B.; Jordan, P. W.; and Pylkkänen, L. 1997. The COCONUT project: Dialogue annotation manual. Available from http://www.isp.pitt.edu/~intgen/research-papers.

Fox, B. 1987. Interactional reconstruction in real-time language processing. *Cognitive Science* 11:365–388.

Green, N. L., and Carberry, S. 1994. Conversational implicatures in indirect replies. In Thompson, H. S., ed., *Proceedings of the Thirtieth Annual Meeting of the Association for Computational Linguistics*, 64–71. San Francisco: Association for Computational Linguistics.

Hancher, M. 1979. The classification of co-operative illocutionary acts. *Language in Society* 8(1):1–14.

Hobbs, J.; Stickel, M.; Appelt, D.; and Martin, P. 1993. Interpretation as abduction. *Artificial Intelligence* 63(1–2):69–142.

Krippendorff, K. 1980. *Content Analysis: an Introduction to its Methodology.* Beverly Hills: Sage Publications.

Lambert, L., and Carberry, S. 1991. A Tripartite Plan-Based Model of Dialogue. In *ACL91, Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, 47–54.

Lambert, L., and Carberry, S. 1992. Modeling Negotiation Subdialogues. In *ACL92, Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics*, 193–200.

Litman, D., and Allen, J. 1990. Discourse Processing and Commonsense Plans. In Cohen, P.; Morgan, J.; and Pollack, M., eds., *Intentions in Communication.* MIT Press.

Lochbaum, K. E. 1994. *Using Collaborative Plans to Model the Intentional Structure of Discourse.* Ph.d. dissertation, computer science department, Harvard University, Cambridge, MA.

McCarthy, J., and Buvač, S. 1995. Formalizing context (expanded notes). Available from http://www-formal.stanford.edu/buvac.

Perrault, C. R., and Allen, J. F. 1980. A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics* 6:167—182.

Qu, Y.; Di Eugenio, B.; Lavie, A.; Levin, L.; and Penstein Rosé, C. 1997. Minimizing cumulative error in discourse context. In Maier, E.; Mast, M.; and LuperFoy, S., eds., *Dialogue Processing in Spoken Language Systems*, Lecture Notes in Artificial Intelligence. Springer Verlag.

Reithinger, N., and Maier, E. 1995. Utilizing statistical dialogue act processing in Verbmobil. In *ACL95, Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics.*

Searle, J. R. 1975. A taxonomy of illocutionary acts. In Gunderson, K., ed., *Language, Mind, and Knowledge. Minnesota Studies in the Philosophy of Science, Vol. 7.* Minneapolis, Minnesota: University of Minnesota Press. 344–369.

Siskind, J. M., and McAllester, D. A. 1993. Nondeterministic lisp as a substrate for constraint logic programming. In *Proceedings of AAAI-93.* CA: AAAI Press.

Stickel, M. 1988. A prolog-like inference system for computing minimum-cost abductive explanations in natural-language interpretation. Technical Report 451, SRI International, 333 Ravenswood Ave., Menlo Park, California.

Strawson, P. F. 1964. Intention and convention in speech acts. *The Philosophical Review* 59:439–460.

Thomason, R., and Hobbs, J. 1997. Interrelating interpretation and generation in an abductive framework. In *Proceedings of AAAI Fall Symposium on Communicative Action in Humans and Machines.* American Association for Artificial Intelligence.

Thomason, R., and Moore, J. 1995. Discourse context. In Buvač, S., ed., *Formalizing Context*, 102–109. Menlo Park, California: American Association for Artificial Intelligence.

Tuomela, R. 1995. *The Importance of Us.* Stanford University Press.

Walker, M. A. 1993. *Informational Redundancy and Resource Bounds in Dialogue.* Ph.D. Dissertation, University of Pennsylvania.

Whittaker, S.; Geelhoed, E.; and Robinson, E. 1993. Shared workspaces: How do they work and when are they useful? *IJMMS* 39:813–842.